REC'D PCT/PTO 14 MAR 2005

PCT/DK 03/00594

REC'D 31 OCT 2003
WIPO PCT

Kongeriget Danmark

Patent application No.: PA 2002 01351

Date of filing: 13 September 2002

Applicant: (Name and address)
Hvidover Hospital
Kettegård Allé 30
DK-2650 Hvidovre
Denmark

Title: Method of rapid detection of mutations and nucleotide polymorphism using chemometrics

IPC: -

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.

BEST AVAILABLE COPY

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patent- og Varemærkestyrelsen Kongeriget Danmark

Patent- og Varemærkestyrelsen
Økonomi- og Erhvervsministeriet

13 October 2003

Pia Høybye

Pia Høybye-Olsen

Modtaget
13 SEP. 2002
PVS

Method of rapid detection of mutations and nucleotide polymorphism using chemometrics.

Field of invention

The present invention relates to the field of molecular biology and more specifically to methods and kits for detection of hybrid polynucleotides between a target and a labelled probe polynucleotide. The methods and kits may be used for the determination of mutations and polymorphisms in samples containing target polynucleotides. All patent and non-patent references cited in the application are hereby incorporated by reference in their entirety.

Background of invention

Detection of hybridised double-stranded polynucleotides

Methods presently available for detecting nucleic acid sequences involve a relatively large number of steps and are laborious to perform. Some of the well-established methods are described for example in Hybridisation, by B. D. Hames and S. J. Higgins (Eds) IRL Press 1985 and J. A. Matthews et al, Analytical Biochemistry, 169, 1988, 1-25 (Academic Press). Typically the DNA to be assayed is denatured and bound to a solid support, e.g. beads, nitrocellulose or nylon, and then incubated with a probe sequence, complementary to the target, which contains a label, for example ^{32}P, or one half of an affinity pair, for example biotin. After incubation and washing of the solid phase, it is then developed to give a signal, by autoradiography for ^{32}P, or in the case of biotin, typically by the addition of an avidin-enzyme conjugate, further washing and then substrate. Also sandwich assays are known, whereby a capture nucleic acid probe, of sequence complementary to the target, is bound to a convenient solid support. This is then allowed to hybridise to the target sequence. A labelled probe is then allowed to hybridise to another part of the target sequence, and after washing steps, the label is developed as appropriate. Alternatively, the "sandwich" is formed in solution with two probes, one containing the label and the other containing one half of an affinity pair. After hybridisation, the solution is then contacted with the solid phase to which the other half of the affinity pair is bound. Washing and signal development then takes place. All these methods

and their variants require numerous handling steps, incubations, washings, and the like which result in processes which are both laborious and time-consuming.

Nucleic acid sequences may be amplified prior to hybridisation to increase the number of target nucleic acids. Methods for detecting amplified nucleic acid sequences, for example DNA sequences generated by PCR, are as outlined above. Alternatively, they can be analysed by gel electrophoresis, and the existence of a band of a given molecular weight is taken as evidence of the presence of the sequence in the original DNA sample, as defined by the specific primers used in the amplification reaction. Again such methods are relatively time-consuming and laborious.

Many hybridisation assays include some sort of signal amplification, which may comprise binding of more labelled probe to the same target nucleic acid or the use of an enzyme based system, wherein the amount of signal produced per probe can be increased by letting the enzymes convert more substrate to detectable product: A further example of signal amplification involves the use of nucleotide amplifiers as described in US 5,635,352 and US 5,124,246 (assigned to Chiron).

Common for all the known methods is that it is always necessary to perform some sort of washing to remove any unhybridised probe, which will give rise to the same signal as hybridised probes. In many cases, it is necessary to perform a whole series of washing operations to reduce background or false positive signal. These hybridisation and washing steps need to be performed by a technician and/or a robot programmed to perform the routine steps thus contributing significantly to the cost of the assays.

Signal detection in most of the known methods is relatively simple in that only one signal is determined and this is done at one wavelength only. Most often, the determination is a qualitative determination of whether there is a significant signal or not. There are also methods available for the quantitative detection of hybridisation between target and probe. These methods may be used to determine the amount of target nucleotides in the sample but are likewise restricted to data detected at one wavelength.

mutations can be single nucleotide substitutions, deletions, insertions or rearrangements. As an consequence of the Human Genome Project interest has been focussed particularly on single nucleotide polymorphisms (SNP) for various reasons including their usage in genome scanning for diseases-related genes. More than 2 million SNP's has been detected, some but not all of which are directly related to diseases.

Often more than one polymorphism, in particular SNP, are present in the same gene. The significance of this varies, but in coding regions, i.e. regions of the gene coding for the proteins, this is of particular interest: In diploid organisms as human it is of importance on which of the two alleles the polymorphisms are located, as this will determine of 0, 1 or several changes in the same protein is present and consequently influence the function of the protein differently.

A number of methods exist that can be used to detect SNPs, mutations and other polymorphisms. Detailed description of useful methods may be found in Ausubel et al. Current protocols in molecular biology, (2000) John Wiley and Sons, Inc., N.Y. and in Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Among the more important methods are: DNA sequencing (Sanger et al. (1977) Nature, 265: 678-695; Maxam and Gilbert (1977) Proc Natl Acad Sci USA, 74: 560-564); the single strand conformation polymorphism (SSCP) method (Orita et al. (1989) Genomics, 5:874-879); the denaturing gradient gel electrophoresis (DGGE) method (Fisher and Lerman. (1993) Proc. Natl. Acad. Sci. U.S.A., 80:1579-1583) and later improvements of the SSCE and the DGGE techniques, such as the dideoxy fingerprinting (Sarkar et al., Genomics, 13:441-443 (1992), restriction endonuclease fingerprinting (Liu and Sommer, Biotechniques, 18:470-477 (1995)) and constant denaturing gel electrophoresis (CDGE), (Hovig et al., Mut. Res., 262:63-71 (1991)). However, all of these methods are labour-intensive, and involves fairly complex handling.

Mutations can be detected by specific hybridisation of oligonucleotides to the nucleic acid to be analysed.

spectral data must be in a format so that they can be exposed to multivariate analysis.

For the purposes of the present invention, performance of mass spectrometry is also regarded as recording a spectrum, although in this case no electromagnetic radiation is emitted or recorded. Also in this case no detectable label is used.

Label: for the purpose of the present invention a detectable label refers to a chemical moiety capable of emitting, absorbing or scattering electromagnetic radiation. More preferably, a label refers to chemical moieties capable of emitting, absorbing, or scattering phosphorescence, luminescence, or fluorescence.

Hybrid polynucleotide: a hybrid between two polynucleotides, wherein the association between the two polynucleotides is stronger than the association between one of the polynucleotides and water.

Summary of invention

Accordingly, the methods according to the present invention provide simple analytical techniques compared to the prior art methods for determination of hybrid polynucleotides. This will render the analytical techniques according the present invention available for laboratories without special molecular biological expertise. The methods are simple and cheap both with regard to equipment and use.

The invention will also solve some of the problems of the research into diseases having a genetic factor. The problem is that many of the important diseases, such as diabetes, artheresclerotic and psychiatric diseases are of polygenic nature meaning that the interaction between several genes leads to a predisposition or directly leads to these diseases. Elucidation of the genetic component of these diseases requires a large number of patients and a significant analytical effort.. The present invention partly reduces the costs associated with these analyses and partly accelerates the investigations.

The analytical method is fast, easy to use and can be completely automated. It is easy to include various control steps in the method and thereby ensure that the detected results can be verified.

In a further aspect the invention relates to a method for detecting a hybrid between a target polynucleotide and a probe oligonucleotide comprising

providing a sample comprising at least one target polynucleotide,

providing at least one polynucleotide probe at least partly complementary to the target polynucleotide,

wherein the polynucleotide probe comprises at least one detectable label,

forming a hybrid polynucleotide comprising at least one target polynucleotide and at least one polynucleotide probe,

recording spectral data representing electromagnetic radiation, said spectral data being recorded from an environment comprising the hybrid polynucleotide, and

treating the spectral data using multivariate analysis to correlate the recorded data to the presence or absence of at least one hybrid.

By using multivariate analysis of spectral data, more information can be gathered from one single hybridisation procedure. Moreover the analysis allows extraction of information which is otherwise "hidden" when looking at the raw data. For example as described in the detailed description part of the present invention, it is possible to group spectra from different signal sources (probe alone, probe-target with 100 % complementarity, probe-target with mismatch, two probes- one traget, etc) and in this way place an unknown spectrum in the correct group. This allows the use of two or more hybridisation reactions to be carried out at the same time; because the data originating from each hybrid can be distinguished by the multivariate analysis. The method also allows determination of the presence or absence of a hybrid without washing away unbound probe, because the multivariate analysis can distinguish the two after appropriate calibration.

instructions enabling correlation of spectral data recorded from a hybrid polynucleotide between said at least one oligonucleotide probe and said target polynucleotide to the presence or absence of said mutation or polymorphism using multivariate analysis.

The instructions may be in the form of calibration data relating to the specific assay to which the kit relates, so that the user of the kit does not have to carry out extensive analyses to determine the grouping of spectra in multivariate analysis.

In a further aspect the invention relates to a system for detection of a hybrid polynucleotide comprising

at least one oligonucleotide probe being at least partly complementary to a target polynucleotide, the probe comprising a detectable label,
a sample chamber from which electromagnetic radiation can be recorded,
a source of spectrally resolved electromagnetic radiation,
means for sensing and recording a spectrum of electromagnetic radiation from the sample chamber,
a computer unit for storing spectral data of electromagnetic radiation and having instructions to treat the recorded spectral data using multivariate analysis.

The system is adapted for performance of the methods according to the present invention and provides a system which allows high throughput screening of samples.

Description of Drawings

Figure 1. The underlying idea of PCA modelling is to replace a complex multidimensional data set by a simpler version involving fewer dimensions, but still fitting the original data closely enough to be considered a good approximation. A 3-D data swarm is shown. The new axes (the principal components PC1 and PC2) are placed in the directions of the largest variances. In this example 3 dimensions are reduced to 2.

absence of wildtype target DNA (W1). The prefix (1-3) gives the replicate number. PCA scores plot from experiment 2.

Figure 9. PCA scores plot based on excitation spectra obtained from 30 °C hybridisations of Cy5 labelled oligonucleotides (A, B, C or D) in the presence of wildtype target DNA (W1). Data points are the same as for Figure 8 but spectra of fluorescent oligonucleotides without targets have been left out. The prefix (1-3) gives the replicate number. PCA score plot from experiment 2.

Figure 10. PCA scores plot based on excitation spectra obtained from 30 °C hybridisations of Cy5 labelled oligonucleotide A in the presence or absence of wildtype target DNA (W2), mutant target (M) or non-specific target (Z). The prefix (1-3) gives the replicate number. PCA scores plot from experiment 3.

Figure 11. PCA scores plot based on emission spectra obtained from 30 °C hybridisations of Cy5 labelled oligonucleotide A in the presence or absence of wildtype target DNA (W2), mutant target (M) or non-specific target (Z). The prefix (1-3) gives the replicate number. PCA scores plot from experiment 3.

Figure 12. PCA scores plot based on excitation spectra obtained from 30 °C hybridisations of Cy5 labelled oligonucleotide A in the presence or absence of wildtype target DNA (A-W2), mutant target (A-M) or both DNA targets simultaneously (A-W2-M, note concentrations: 0,125 or 0,25 µM). The measurement A-W2-M-2_2 is not included while it was considered as an outlier. The prefix (-1 or -2) represent two different concentrations of the targets, and the prefix (1-3) give the replicate number. Data are from experiment 4. An outlier has been removed from the data analysis.

Figure 13. PCA scores plot based on emission spectra obtained from 30 °C hybridisations of Cy5 labelled oligonucleotide A in the presence or absence of wildtype target DNA (A-W2), mutant target (A-M) or both DNA targets simultaneously (A-W2-M). The prefix (-1 or-2) represent two different concentrations of the targets, and the prefix (1-3) give the replicate number. Data are from experiment 4.

Figure 17/18. The buffer and target alone do not contribute significantly to the spectrum of Figures 17/18.

In probes A and B, the mutant nucleotide is placed terminally and the label is linked directly to the mutant nucleotide. On the right is shown similar spectra of samples with both the target polynucleotide (W1 as defined in Figure 4) and the two probes (A-wildtype and B-mutant). Probe A which is 100 % complementary to a sequence in the target polynucleotide interacts differently with this than probe B, which differs in one position. Through the interaction between target and probe, the recorded spectrum is changed slightly. In this particular case the difference is just visible in the spectrum, but in order to make full use of the information in the spectra multivariate analysis is used. These changes can be analysed using multivariate analysis. It can also be seen that the spectrum of probe B, which differs from probe A at one position (see Figure 4) is not changed to the same degree by interaction with the target polynucleotide.

In Figure 18 a further example is illustrated. The probes (C and D) differ from probes A and B in that the mutant nucleotide is placed centrally in the probe (see Figure 4). The spectra to the right of Figure 18 are recorded from samples with both the probe and the target polynucleotide, which is 100 % complementary to probe C and differs from probe D at one (terminal) position. Again the association between target and probe changes the spectrum so that the presence of a target wildtype can be distinguished form a target mutant polynucleotide.

Data analysis

The art of extracting chemically relevant information from data produced in chemical experiments is given the name "Chemometrics" in analogy with biometrics and econometrics. Chemometrics is heavily dependent on the use of different kinds of mathematical models - high information models, ad-hoc models, and analogy models. The task demands knowledge of statistics, numerical analysis, operation analysis, etc., and applied mathematics. However, as in all applied branches of science, the difficult and interesting problems are defined by the applications.

data set includes noise and since this noise is not used to describe the differences between samples there is a limit to how much variation one can describe.

Figure 1 visualises how the 2 first principal components (PC1 and PC2) are found in a 3 dimensional plot. Only data of 3 variables are available, so the linear behaviour is immediately recognised by plotting the objects in the 3-D variable space. When more variables are present (like in spectroscopy where each row of a data matrix is a spectrum of perhaps several hundred of wavelengths), this procedure is of course not feasible. Identification of this type of linear behaviour in a space with several hundred dimensions of course cannot be achieved by visual inspection. Here PCA, with its powerful projection characteristics, helps to discover the hidden data structures.

In practice, a PCA involves 3 steps:

- Choice of pre-processing method, i.e. data pre-treatment method
- Running the PCA algorithm, selection of number of PCs, evaluation of the model
- Interpretation of plots (loading and score plots)

Pre-processing is used to ensure that the raw data have a distribution that is optimal for the analysis. Background effects, measurements with various units, different variance of the variables, etc. makes it difficult to extract the meaningful information. Pre-processing reduces the noise introduced by such effects.

Pre-processing may contain the operations of centering, weighting and transformations. In the examples transformations have not been used. Transformations include logarithming, smoothing, deriving, normalisation, scatter correction.

In (mean-) centering the average of the data of each variable is subtracted. This ensures that all results will be interpreted in terms of variation around the mean. This centering method has been performed in all the analyses/models in the examples.

is, these variables are most prominent in explaining the differences between samples.

On the other hand, it also means that variables that lie very close to zero in the plot (PARAM.1) are of no significance to the separation of the samples, i.e. these could be spared. It is thus possible to identify the variables that are dominating to the samples that have been tested.

The mutual location of the variables in the loading plot, reveals which variables provide the same information about the samples. Variables that lie on the same straight line through zero in the plot hang together: If they lie on the same side of zero (PARAM.5 and PARAM.6) then there is a positive correlation, i.e. when one is high then the other is too, but if they lie on both sides of zero (PARAM.5 and PARAM.7), then there is a negative correlation, i.e. one is high and the other is low. Variables that lie like this supplement only each other. They do not contribute additionally to the information concerning the differences between the samples. If, for example, three variables lie on the same spot in the plot, one can be content with the one and still obtain the same information.

By comparing the scores and loading plots it can finally be determined which of the variables that determine the grouping of different samples or sample clusters. When the samples lie on the "same spot" in the score plot (relatively, for example in the right corner of the plot) as one or several variables in the corresponding loading plot, the samples have relatively high values for these variables. One can, in other words, ascertain which variable is the dominating one to the different samples. In the given example the variables, PARAM.5 and PARAM.6, are essential to sample2 and sample6. Likewise, the variable, PARAM.4, is essential to sample3

Outliers

The definition of an outlier is an observation (outlying sample) or variable (outlying variable) which is abnormal compared to the major part of the data.

Extreme points are not necessarily outliers; outliers are points that apparently do not belong to the same population as the others, or that are badly described by a model.

cluster analysis, but rather as a tool to develop and verify causal interrelationships between items, events etc. These two groups of techniques are therefore complementary rather competitive. In the present context PCA is used for basic data-analyses and reduction. In further analysis of the spectra, in particular when two or more labels are used, the above-mentioned techniques will be applicable in data-treatment and presentation. E.g if two labels are used, spectre from two labels will be obtained, which preferable should be analysed separately. Clustering of the spectre, analysis of interaction between the spectre of each of the labels concomitantly obtained and presenting data in a comprehensible and interpretable way will use elements of all the above mentioned techniques.

Spectroscopy

Spectroscopic techniques form the largest and most important single group of techniques used in analytical chemistry, and provide a wide ranges of quantitative and qualitative information. All spectroscopic techniques depend on the emission or absorption of electromagnetic radiation characteristic of certain energy changes within an atomic or molecular system. The energy changes are associated with a complex series of discrete or quantised energy levels in which atoms and molecules are considered to exist.

The relevant spectroscopic techniques included but are not limited to:

Table 1

arc/spark spectrometry
spectrography
plasma emission spectrometry
flame photometry
X-ray fluorescence spectrometry
atomic fluorescence spectrometry
atomic absorption spectrometry
γ-spectrometry
ultraviolet spectrometry
visible spectrometry
infrared spectrometry

The method according to any of the preceding claims, wherein the spectral data recorded comprises a fluorescence spectrum between 180 and 950 nm.

A fluorescence spectrum may be an excitation spectrum or an emission spectrum, or both, depending on the level of reduction of variance not related to the interaction of target and non-target. To the extent that these non-interactive variances can be eliminated e.g. by clustering the spectre both emission and excitation spectre alone or in combination can be of analytical value.

In order to better analyse the difference between a spectrum from an unbound and a bound probe, the method preferably further comprises recording of spectral data from the polynucleotide probe alone. After having established this difference, it is expected that it is not always necessary to record this spectrum, although it may in some instances be useful as a calibration for day-to-day variation.

The method may further comprise recording spectral data from the hybrid polynucleotide and from a polynucleotide probe alone and/or, from a non-hybridising polynucleotide probe contacted by the traget polynucleotide, and/or from a polynucleotide probe contacted with a non-hybridising polynucleotide sequence. All of these spectra, which may be regarded as "controls" serve the purpose of calibrating the method.

Mass spectrometry is a technique for characterising molecules according to the manner in which they fragment when bombarded with high-energy electrons, and for elemental analysis at trace levels. It is not strictly speaking a spectrometric method as electromagnetic radiation is neither absorbed nor emitted. However, the data obtained are in a spectral form in that the relative abundance of mass fragments from a sample is recorded as a series of lines or peaks. The bombardment process produces many fragments carrying a charge, and this facilitates their separation and detection by electrical and magnetic means. Spectra must be recorded under conditions of high vacuum (10^{-4} to 10^{-6} Nm^{-2}) to prevent loss of the charged fragments by collision with molecules of atmospheric gases or swamping of the sample spectrum.

Cy7 (trademarks for Biological Detection Systems, Inc.), fluorescein, acridin, acridin orange, Hoechst 33258, Rhodamine, furthermore: Rhodamine Green, Tetramethylrhodamine, Texas Red, Cascade Blue, Oregon Green, Alexa Fluor (trademarks for Molecular Probes, Inc.), 7-nitrobenzo-2-oxa-1-diazole (NBD), pyrene and Europium, Ruthenium, Samarium, and other rare earth metals, bimane, ethidium, europium (III) citrate, La Jolla blue, methylcoumarin, nitrobenzofuran, pyrene butyrate, rhodamine, and terbium chelate. More specialised fluorochromes are listed in Table 2 along with their suppliers.

TABLE 2

FLUORESCENT LABELS

Fluorochrome	Vendor	Absorption Maximum	Emission Maximum
Bodipy 493/503	Molecular Probes	493	503
Cy2	BDS	489	505
Bodipy FL	Molecular Probes	508	516
FTC	Molecular Probes	494	518
FluorX	BDS	494	520
FAM	Perkin-Elmer	495	535
Carboxyrhodamine	Molecular Probes	519	543
EITC	Molecular Probes	522	543
Bodipy 530/550	Molecular Probes	530	550
JOE	Perkin-Elmer	525	557
HEX	Perkin-Elmer	529	560
Bodipy 542/563	Molecular Probes	542	563
Cy3	BDS	552	565
TRITC	Molecular Probes	547	572
LRB	Molecular Probes	556	576
Bodipy LMR	Molecular Probes	545	577
Tamra	Perkin-Elmer	552	580
Bodipy 576/589	Molecular Probes	576	589
Bodipy 581/591	Molecular Probes	581	591
Cy3.5	BDS	581	596
XRITC	Molecular Probes	570	596
ROX	Perkin-Elmer	550	610
Texas Red	Molecular Probes	589	615

order to quench parts of the spectrum, which only contributes to noise. This would be most appropriate when non-relevant variance (noise) cannot be eliminated by multivariate analysis. In some instances quenchers may act as modifiers to alter the general spectral characterstics.

Examples of quenchers depend on the specific part(s) of the spectrum to be quenched. The quenching abilities of a number of different compounds are known in the art. One particular example is the TAMRA quencher.

Immobilisation to solid surface

In some embodiments of the present invention it is preferable that one or more probes or targets polynucleotide is immobilised to a solid surface. The nature of the means for immobilisation and of the nature of the solid support is a matter of choice. Numerous suitable supports and methods of attaching nucleotides to them are well known in the art and widely described in the literature. Thus for example, supports in the form of microtiter wells, tubes, dipsticks, particles, fibers or capillaries may be used, made for example from agarose, cellulose, alginate, teflon, latex or polystyrene. Conveniently, the support may comprise magnetic particles, which permits the ready separation of immobilised material by magnetic aggregation.

The solid support may carry functional groups such as hydroxyl, carboxyl, aldehyde or amino groups for the attachment of nucleotides. These may in general be provided by treating the support to provide a surface coating of a polymer carrying one of such functional groups, e.g. polyurethane together with a polyglycol to provide hydroxyl groups, or a cellulose derivative to provide hydroxyl groups, a polymer or copolymer of acrylic acid or methacrylic acid to provide carboxyl groups or an amino alkylated polymer to provide amino groups. US 4,654,267 describes the introduction of many such surface coatings.

Alternatively, the support may carry one member of an "affinity pair", such as avidin, while the polynucleotide is conjugated to the other member of the affinity pair *in casu* biotin. Representative specific binding affinity pairs are shown in Table 4.

TABLE 4. Representative Specific Binding Pairs

dependent amplification methods results in labelled DNA fragments which can be immobilised on specialised surfaces. For instance SH-modified DNA may be immobilised on a gold surface (Steel et al. (2000) Biophys J 79:975-81) likewise 5'-phosphorylated DNA or 5'-aminated DNA may be immobilised by reaction with activated surfaces (Oroskar et al. (1996) Clin Chem 42:1547-55; Sjoroos et al (2001) Clin Chem 47:498-504).

During synthesis oligonucleotides may also be labelled or coupled to photoreactive groups. Acetophenone, benzophenone, anthraquinone, anthrone or anthrone-like modified DNA can for instance be activated by exposure to UV light and immobilised on a wide range of surfaces as described in European and US patents: EP0820483, US6033784 and US5858653. Also photoreactive psoralens, coumarins, benzofurans and indols have been used for immobilisation of nucleic acids. An extensive discussion of immobilisation of nucleic acids can be found in WO 85/04674.

Target polyncleotides

The methods, kits and systems of the present invention can be used with any kind of polynucleotide, which can be subjected to a hybridisation assay. One source of target polynucleotide are RNA polyncleotides, such as wherein the target polynucleotide comprises RNA, such as mRNA and/or rRNA and or tRNA.

Among the RNAs the rRNAs including 5S, 5.5-5.8S, 16S, 18S, 23S, 25-28S rRNA can be used for identifying the taxon from which the RNA was isolated, since conserved sequences can be found for a number of taxons and for a number of species. rRNA may for example be used for diagnosing an infectious disease caused by microorganisms, or for determining the amount and nature of contamination in food, feed and various water-supplies.

Another source of target polynucleotides are DNA. DNA may be used for determination of mutations and/or polymorphisms, but also in genotyping individuals or for determining the taxon, for forensic usage or for linkage studies. The different sources of DNA include genomic DNA, organelle DNA, mitochondrial DNA, chloroplast DNA, cDNA, and environmental DNA.

bases, from 35000 to 40000 bases, from 40000 to 45000 bases, from 45000 to 50000 bases, more than 50000 bases.

When referring to the length of the target polynucleotide, the whole length of the molecule is intended. During hybridisation, the probe only hybridises to one or more relatively short sequence(s) of the target polynucleotides. These sub-sequences to which the probe hybridises are termed the range of overlap between target and probe. This length of the overlap between the probe and target polynucleotide may be as short as at least 5 nucleotides. However more specific hybridisation is obtained by increasing the length of the overlap. Therefor, more preferably the overlap is at least 6 nucleotides, such as at least 7 nucleotides, for example 8 nucleotides, such as at least 9 nucleotides, for example at least 10 nucleotides, such as at least 15 nucleotides, for example at least 20 nucleotides, for example at least 25 nucleotides, such as at least 50 nucleotides, for example at least 100 nucleotides.

Extraction of target nucleic acids

Extraction of target nucleic acids can be performed using methods known to those skilled in the art (Joseph Sambrook & David W Russell (2001) "Molecular cloning: a laboratory manual", Cold Spring Harbor Laboratory Press, New York, USA). When selecting extraction protocols among the numerous available protocols, it is preferable to select protocols where the buffer ingredients do not interfere with the recordation of spectral data. However, it is also possible to precipitate the extracted target nucleic acid to get rid of components, which interfere with the spectrum.

Probes

The probes used in the method according to the present invention may be made from any kind of nucleotide monomer or any combination of the known types of monomers. Thus the probe may comprise at least one RNA monomer, and/or comprises at least one DNA monomer, and/or at least one PNA monomer, and/or at least one methylated monomer, and/or at least one labelled monomer, and/or at least one LNA monomer.

Generally, it can be said that those probes are preferred, where the label is positioned as close as possible to the polymorphic site to maximise the spectral difference. The distance from the label to the polymorphic site may be 1 nucleotide, 2 nucleotides, 3 nucleotides, 5 nucleotides, 10 nucleotides or more. Among these, the most preferred are those where the nucleotide complementary to the polymorphic site is in a terminal position.

apolipoprotein B mutations related to atherosclerosis, wherein the probe may comprise a sequence from any of SEQ ID NO 1 to 4.

apolipoprotein E polymorphism (apoE2, E3 and E4) related to neurological diseases, wherein the probe may comprise a sequence from any of SEQ ID NO 5 to 8.

human muscle glycogen synthase polymorphism related to diabetes mellitus, wherein the probe may comprise a sequence from any of SEQ ID NO 9 to 10.

methylene tetrahydrofolate reductase polymorphism related to osteoporose, wherein the probe may comprise a sequence from any of SEQ ID NO 13 to 14.

Dnase1 mutations related to rheumatological diseases, wherein the probe may comprise a sequence from any of SEQ ID NO 11 to 12.

BRCA1 gene or in the BRCA2 gene related to breast or ovarian cancer, wherein the probe may comprise a sequence from any of SEQ ID NO 27 to 30.

Mismatch repair gene mutations related to cancer, wherein the probe comprises a sequence from any of SEQ ID NO 15 to 16.

The probe may be selective for a mutation in a promoter sequence, or the probe may be selective for a mutation in a coding sequence, including introns and exons.

Apart from mutations, the probes may be used to diagnose the presence/absence and/or nature of a microbial infection, wherein the probe is selective for a microbial

spectral data from a solution comprising both the hybrid polynucleotide and unhybridised probe. The spectrum form the complex solution is changed by the mere presence of the hybrid.

The ratio of target to probe in the hybridisation solution may range from 1:0.1 to 1:10, such as 1:0.2, 1:0.5, 1:0.75, 1:1; 1:2, 1:4, 1:5, 1:7, 1:8, or 1:10.

However, it is likewise possible to record spectral data from hybrid polynucleotides bound to a solid support. The solid support may comprises a solid surface capable of immobilising a capture probe, a capture probe capable of immobilising the target polynucleotide, and a labelled detection probe capable of hybridising to the immobilised target polynucleotide. The capture probe may be is immobilised a priori to the solid surface or the capture probe may be hybridised to the target before immobilisation on a solid support. In a separate embodiment the the capture probe(s) is/are (an) allele specific probe(s).

In a special layout the solid support is a disposable or reusable device such as but not exclusively a flow-through system. Such a flow through system may comprise immobilised capture probes to capture the target, which can then be labelled by hybridisation with a label probe. However, it is also possible to label the capture probe directly, and obviate the need for a separate label probe.

Amplification

The target polynucleotides may be amplified by one of many methods. One of the best known and widely used amplification methods is the polymerase chain reaction (referred to as PCR) which is described in detail in US 4,683,195, US 4,683,202 and US 4,800,159, however other methods such as LCR (Ligase Chain Reaction, see Genomics (1989) 4:560-569), NASBA (Nucleic Acid Sequence-Based Amplification, see *PCR Methods Appl* (1995) 4, S177-S184), strand displacement amplification (*Current Opinion in Biotechnology* (2001) 12:21-27) or rolling circle amplification (*Current Opinion in Biotechnology* (2001) 12:21-27) can be applied. Genuine amplification in e.g. bacteria or yeast can also be used to amplify nucleotide sequences as can non-PCR methods, such as T7-polymerase.

backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone (*Science* (1991) 254: 1497-1500).

In general, the hybridisation conditions are chosen so that the formation of a hybrid polynucleotide takes place under conditions of

optimal or suboptimal stringency providing sufficient stable complexes for discriminatory signal detection,

any composition of buffers optimising discriminatory signal detection,

any form and concentrations of one or more salts optimising discriminatory signal detection,

any additives including but not limited to stabilisers and/or quenchers optimising discriminatory signal detection,

temperature range for hybridisation specific for any specific combination of analyte and probe optimising discriminatory signal detection,

any range of time of hybridisation necessary to optimise discriminatory signal detection.

As evidenced by the appended examples the hybridisation temperature can be varied within the normal limits. In other words, the formation of a hybrid may be performed at a temperature between 10 and 90°C such as 10 to 20 °C, 20-30 °C, 30 to 40 °C, 40 to 50 °C, 50 to 60 °C, 60 to 70 °C, 70 to 80°C, or 80 to 90°C.

The buffer used for the hybridisation is conveniently a PCR buffer, which preferably is non-fluorescent, and/or which stabilises the spectrum of electromagnetic radiation, and which allows hybridisation.

Determination of haplotypes

The methods of the present invention may likewise be used for the determination of haplotypes. The haplotype is the set of alleles borne on one of a pair of homologous chromosomes. Often the particular combination of alleles in a defined region of some chromosome, e.g. the locus of the major histocompatability complex (MHC), is referred to as the haplotype of that locus. The central dogma of modern molecular genetics teaches that it is the haplotype of the coding part of a gene that determines

molecules. This difference between intra- and inter-molecular hybridisation is in particular significant when hybridisation is performed at low concentrations of nucleic acid and results in a significant difference even between the hybridisation of the bifunctional probe to one or two amplified fragments, the hybridisation to one fragment being the most favourable. Under stringent conditions employed in the hybridisation only probes which are completely complementary to the sequences comprising both studied polymorphisms will form stable hybrids.

Following the stringent hybridisation, the haplotype can be determined by recording a fluorescence spectrum from the different oligonucleotides hybridised to the amplified DNA without the need for separation of hybrid from .

Further details on how to perform determination of haplotypes can be found in PCT/DK02/00552 (HVIDOVRE HOSPITAL), which is hereby incorporated by reference in its entirety. ,

Kits for detection of mutations/polymorphisms

The invention also relates to a kit for detection of mutations or polymorphism comprising

at least one oligonucleotide probe capable of hybridising to a preselected region of a target polynucleotide, the polynucleotide probe further comprising at least one detectable label,

instructions enabling correlation of spectral data recorded from a hybrid polynucleotide between said at least one oligonucleotide probe and said target polynucleotide to the presence or absence of said mutation or polymorphism using multivariate analysis.

The kit is assembled and sold with probes and instructions enabling the performance of one of a number of different assays as outlined above. The instructions can be regarded as calibration data, which enable the user to perform the hybridisation assay without first determining the difference between a spectrum from an unbound an a bound probe. The instructions are generally rather complex and may conveniently be stored on a data storage medium, which is provided together with the probes, and optional buffers. Instructions may be in the form of

detectably different labels, which preferably do not interfere with each other's spectum. In this way hybridisation to the two probes can be detected with one scan over a given wavelength range.

Systems for detecting a hybrid polynucleotide

In a further aspect the invention relates to a system for detection of a hybrid polynucleotide comprising

at least one oligonucleotide probe being at least partly complementary to a target polynucleotide, the probe comprising a detectable label,

a sample chamber from which electromagnetic radiation can be recorded,

a source of spectrally resolved electromagnetic radiation,

means for sensing and recording a spectrum of electromagnetic radiation from the sample chamber,

a computer unit for storing spectral data of electromagnetic radiation and having instructions to treat the recorded spectral data using multivariate analysis.

These systems are adapted for high-throughput screening of hybridisation events.

The system may further comprising a computer controlled robot to transfer solutions to the sample chamber.

Preferably the system comprises means to control the temperature of the sample chamber during hybridisation and subsequent recording of spectrum.

As for the kits, the sample chamber may be in the form of a tube with at least one probe linked to the inner surface. The tube may comprise more than one probe linked to more than one spatially separate location and the system comprises means to record a spectrum from each of the spatially separate locations. And the tube may comprise more than one probe, the more than one probe having detectably different labels.

The system preferably is made to accommodate standard laboratory glass and/or plasticware, and may in one preferred embodiment be adapted to accommodate a

Examples

METHODS

Oligonucleotides

The oligonucleotides used for the examples were selected from a region of the Apolipoprotein gene - Apo3611. The oligos are aligned in Figure 4 to show the regions of complementarity.

Unlabelled oligonucleotides

Wildtype 1

W1

5´CTAAGAACCAGAAGATCAGATGGAAAAATGAAGTCCGGATTCATTCTGGGTCTTTCCAGA GCCAGGTCGA

Wildtype 2

W2 5´ACCAGAAGATCAGATGGAAAAATGAAGTCC**G**GATTCATTCTGGGTCTTTC

Wildtype 3

W3

5´GCTAACACTAAGAACCAGAAGATCAGATGGAAAAATGAAGTCC**G**GATTCATTCTGGGTCT TTC

Mutant

M 5´ACCAGAAGATCAGATGGAAAAATGAAGTCC**A**GATTCATTCTGGGTCTTTC

Reverse oligonucleotide

R 3´TGGTCTTCTAGTCTACCTTTTTACTTCAGGCCCAAGTAAGACCCAGAAAG

Y Double stranded target molecule formed by W2 and R.

Non-specific target molecule

Z 5´GTTCACGAGCTCAGCAACCTGTGACCTGAATTCAGTCTGATAAAATCGCA

Cy5 labelled oligonucleotides

A (Cy5-wt-ter)

Cy5-5'-**C**GGACTTCATTTTTC

B (Cy5-mu-ter)

Cy5-5'-**T**GGACTTCATTTTTC

connection to a termostated water bath. Individual oligonucleotide mixtures were kept at the hybridisation temperatures in a termostated water bath. Spectral ranges are provided in Table 5. The excitation spectrum was obtained by keeping the Em_{max} constant, whereas the emission spectrum was obtained by keeping the excitation wavelength at Ex_{max}.

Dye-oligo	Ex_{max} nm	Em_{max} nm	Ex spectrum nm	Em spectrum nm
Cy3-oligo	550	570	500 – 800	500 - 800
Cy5-oligo	649	670	550 – 800	550 - 800
Rhodamine-oligo	570	587	500 – 800	500 - 800

Table 5. Fluorescence properties of the three fluorophores used in the present study.

Experiments

Nine experiments were performed during the investigation.

Experiment 1: Two pairs of Cy5-5' labelled oligonucleotides which were either perfect match or mismatch, respectively, in the 5'-ends or in the middle positions, relative to the target were hybridised to a wildtype target (W1). Spectra were obtained at RT (23 °C).

Experiment 2: Same as Exp 1, but in addition a set of rhodamine labelled oligonucleotides were used. Spectra were obtained at 30 °C after 5 min of hybridisation.

Experiment 3: A Cy5 labelled oligonucleotide (complementary to the wildtype target) was hybridised to the wildtype target (W2), the mutant target (M) or the dummy target (Z), respectively. Spectra were obtained at 30 °C after 5 min of hybridisation.

Experiment 4: Hybridisation of the Cy5 labelled wildtype oligonucleotide to either the wildtype target (W2), the mutant target (M) or both W2 and M simultaneously.

Experiment 5: Hybridisation of two sets of Cy3 and Cy5 labelled oligonucleotides simultaneously (only partial measurements).

The first prefix (1 or 2) denotes the replicate number, whereas the second prefix only represent two fluorescence measurements on that same replicate solution. This implies, that samples with the prefix 1.1 and 1.2 and the prefix 2.1 and 2.2 respectively should be placed closer to each other in the plot – only showing the difference in fluorescence spectra obtained successively in time. The plot shows that this is the case. There is most often a larger variation within replicates than within two fluorescence measurements on the same replicate. In conclusion there are some differences between replicates and double measurements. It should be noted that the hybridisations reactions were not allowed to hybridise for exactly the same time in this first shot experiment.

Conclusion of experiment 1:

Hybridisation changes the fluorescence emission spectra.
The 4 hybridised oligonucleotide combinations (A-W1, B-W1, C-W1 and D-W1) results in different emission fluorescence spectra.

Experiment 2

Two pairs of Cy5-5' labelled oligonucleotides which were either perfect match or mismatch, respectively, in the 5'-ends or in the middle positions, relative to the target were hybridised to a wildtype target (W1). Spectra were obtained at 30 °C.

Figure 6 shows, that when the 4 different oligonucleotides (A, B, C and D) are hybridised to a target DNA (W1) it changes their placement in the scores plot - i.e. the hybridisation changes the obtained fluorescence spectrum (emission).

For all the 4 oligonucleotides the hybridisation to the target DNA moves the samples more to the right in the scores plots, which implies that the score value of PC1 increases.

In Figure 7 it is seen, that the 4 oligonucleotides (A, B, C and D) hybridised to the target DNA (W1) is placed in each group of the scores plot - i.e. the 4 hybridised oligonucleotides gives different fluorescence spectra (emission). However, the variation within each group is just a little smaller than the variation between groups.

Conclusions of experiment 3:

Hybridisation changes the fluorescence spectrum. (both the excitation and the emission spectra).

The 3 hybridised oligonucleotides (A-W2, A-M and A-Z) separates out in different groups.

Experiment 4

Hybridisation of the Cy5 labelled wildtype oligonucleotide to either the wildtype target (W2), the mutant target (M) or both W2 and M simultaneously.

Figure 12 (excitation spectrum) shows that when oligonucleotide A is hybridised to a target DNA (W2, M or W2/M) it changes its placement on the scores plot - i.e. the hybridisation changes the obtained fluorescence spectrum (excitation).

It is observed, that the 2 hybridised oligonucleotides A-W2 and A-M are placed in each group in the scores plot. The hybridised oligonucleotides A-W2-M (W2 and M targets are present simultaneously) are placed in between the A-W2 group and the A-M group. No clear difference between the two W2-M concentrations (corresponding to the prefix –1 or –2 in the A-W2-M samples) was observed, i.e. the two concentrations (0,25 or 0,125 μM of each oligo nucleotide) did not influence on the grouping.

Figure 13 (emission spectrum) gives the same picture as is depicted in Figure 12. Compared to the results observed in Figure 12 which is based on excitation spectra, the results based on the emission spectra indicates a smaller variation within each group in relation to the variation between the groups.

The conclusions of experiment 4 are:

Hybridisation changes the fluorescence spectrum. (both in excitation and in emission).

There is a clear difference between F-A-W3 and F-B-W3 which means that in the presence of a second probe (the Cy3 labelled F) the full match (A-W3) and the mismatch (B-W3) can still be differentiated.

Reference List

Bentzen, J., T. Jorgensen, and M. Fenger. "The effect of six polymorphisms in the apolipoprotein B gene on parameters of lipid metabolism in a Danish population." Clin.Genet. In press (2002).

Bentzen, J. et al. "The influence of the polymorphism in apolipoprotein B codon 2488 on insulin and lipid levels in a Danish twin population." Diabet.Med. 19.1 (2002): 12-18.

Fenger, M. et al. "Impact of the Xba1-polymorphism of the human muscle glycogen synthase gene on parameters of the insulin resistance syndrome in a Danish twin population [In Process Citation]." Diabet.Med. 17.10 (2000): 735-40.

Jarvik, G. P. "Genetic predictors of common disease: apolipoprotein E genotype as a paradigm." Ann.Epidemiol. 7.5 (1997): 357-62.

Peltomaki, P. "Deficient DNA mismatch repair: a common etiologic factor for colon cancer." Hum.Mol.Genet. 10.7 (2001): 735-40.

Yasutomo, K. et al. "Mutation of DNASE1 in people with systemic lupus erythematosus." Nat.Genet. 28.4 (2001): 313-14.

Sequence table

Highlighted residues correspond to the position of polymorphisms in the target nucleic acid sequence.

Apolipoprotein B mutations related to atherosclerosis (Bentzen et al. 12-18;Bentzen, Jorgensen, and Fenger)

Apolipoprotein B 3500 mutation:

SEQ ID NO 19 5'-GGAAGAAGCTTGCTTCTTTGCTGAC (specificcapture-probe for E. Coli-ECA75F 16S)

Experimental oligoes and probes (the oligonucleotides were selected from a region of the apolipoprotein B gene at the codon3611 polymorphism):

SEQ ID NO 20

W1 5-ctaagaaccagaagatcagatggaaaaatgaagtccggattcattctgggtctttccagagccaggtcga

SEQ ID NO 21

W2 5-accagaagatcagatggaaaaatgaagtccggattcattctgggtctttc

SEQ ID NO 22

M 5-ACCAGAAGATCAGATGGAAAAATGAAGTCC**A**GATTCATTCTGGGTCTTTC'

SEQ ID NO 23

A CTTTTTACTTCAGG**C**-5' (Cy5-wt-ter)

SEQ ID NO 24

B CTTTTTACTTCAGG**T**-5' (Cy5-mut-ter)

SEQ ID NO 25

C CTTCAGG**C**CTAAGTA-5' (Cy5-wt-cen)

SEQ ID NO 26

D CTTCAGG**T**CTAAGTA-5' (Cy5-mut-cen)

SEQ ID NO 27 5'-AACACC**C**AGGATCCT (BRCA1 gene codon 1313 wt)

SEQ ID NO 28 5'-AACACC**T**AGGATCCT (BRCA1 gene codon 1313 mut)

SEQ ID NO 29 5'-CTG GAA**C**AGTCTGGG (BRCA1 gene codon 1541 wt)

SEQ ID NO 30 5'-CTG GAA**T**AGTCTGGG (BRCA1 gene codon 1541 mut)

In the following the possible variation in probes is described with reference to the ApoE polymorphism (SEQ ID No 5-8). *Apolipoprotein E polymorphism (apoE2, E3 and E4) related to neurological diseases* (Jarvik 357-62)

The residue corresponding to the polymorphism is located in the 3' end of the probe. The label may be linked to the 3' or the 5' terminal residue.

SEQ ID NO X1 5'-ATGGAGGACGTG**T** apoE codon112-cys

SEQ ID NO X2 5'-ATGGAGGACGTG**C** apoE codon112-arg

SEQ ID NO X3 5'-GACCTGCAGAAG**T** apoE codon 158-cys

Claims

1. A method for identifying a hybrid polynucleotide comprising at least partly complementary nucleotide strands, said method comprising the steps of

   i) providing a sample comprising at least one target polynucleotide,

   ii) providing at least one polynucleotide probe at least partly complementary to the at least one target polynucleotide,

   wherein the polynucleotide probe comprises at least one detectable label,

   iii) forming a hybrid polynucleotide comprising at least one target polynucleotide and at least one polynucleotide probe, and

   iv) recording spectral data from an environment comprising the hybrid polynucleotide,

   v) wherein the spectral data when at least one oligonucleotide probe forms part of the hybrid polynucleotide are different from the spectral data when at least one oligonucleotide probe is not part of the hybrid polynucleotide.

2. A method for detecting a hybrid between a target polynucleotide and a probe oligonucleotide comprising

   i) providing a sample comprising at least one target polynucleotide,

   ii) providing at least one polynucleotide probe at least partly complementary to the target polynucleotide,

   wherein the polynucleotide probe comprises at least one detectable label,

4. The method according to any of the preceding claims 1 to 3, wherein the at least one probe has a length of 6 to 50 nucleotides, preferably 6 to 25 nucleotides, such as 6 to 8 nucleotides, 8-10 nucleotides, 10-12 nucleotides, 12-14 nucleotides, 14-16 nucleotides, 16-18 nucleotides, 18-20 nucleotides, 20-22 nucleotides, or 22-25 nucleotides.

5. The method according to claim 4, wherein the sequence complementarity between target and probe in a range of overlap is at least 50%, more preferably at least 60 %, more preferably at least 70 %, more preferably at least 80 %, more preferably at least 85 %, more preferably at least 90 %, more preferably at least 95 %, more preferably at least 96 %, more preferably at least 98 %, more preferably 100%.

6. The method according to any of the preceding claims, wherein at least one probe comprises at least one RNA monomer.

7. The method according to any of the preceding claims, wherein at least one probe comprises at least one DNA monomer.

8. The method according to any of the preceding claims, wherein at least one probe comprises at least one PNA monomer.

9. The method according to any of the preceding claims, wherein at least one probe comprises at least one methylated monomer.

10. The method according to any of the preceding claims, wherein at least one probe comprises at least one LNA monomer.

11. The method according to any of the preceding claims, wherein at least one probe comprises a mixture of monomers in claims 6-10

12. The method according to any of the preceding claims, comprising using at least two polynucleotide probes capable of hybridising to two different target polynucleotides.

23. The method according to any of the preceding claims, wherein the probe is selective for human muscle glycogen synthase polymorphism related to diabetes mellitus.

24. The method according to claim 23, wherein the probe comprises a sequence from any of SEQ ID NO 9 to 10 or similar probes, where the mutation/polymorphism is placed in the 5' or 3' end.

25. The method according to any of the preceding claims, wherein the probe is selective for methylene tetrahydrofolate reductase polymorphism related to osteoporose.

26. The method according to claim 25, wherein the probe comprises a sequence from any of SEQ ID NO 13 to 14 or similar probes, where the mutation/polymorphism is placed in the 5' or 3' end.

27. The method according to any of the preceding claims, wherein the probe is selective for Dnase1 mutations related to rheumatological diseases.

28. The method according to claim 27, wherein the probe comprises a sequence from any of SEQ ID NO 11 to 12 or similar probes, where the mutation/polymorphism is placed in the 5' or 3' end.

29. The method according to any of the preceding claims, wherein the probe is selective for a mutation in the BRCA1 gene or in the BRCA2 gene.

30. The method according to claim 29, where the probe comprises a nucleotide sequence selected from any of SEQ ID No 27-30 or similar probes, where the mutation/polymorphism is placed in the 5' or 3' end.

31. The method according to any of the preceding claims, wherein the probe is selective for mismatch repair gene mutations related to cancer.

32. The method according to claim 31, wherein the probe comprises a sequence from any of SEQ ID NO 15 to 16.

positive controls (wild-type, mutation, heterozygote), negative control (dummy DNA sequence).

44. The method according to any of the preceding claims, wherein the target polynucleotide comprises chemically or biologically modified nucleic acids.

45. The method according to claim 44, wherein the modification comprises modification of cytosin by bisulphite.

46. The method according to any of claims 38 to 45, wherein the target polynucleotide comprises a mixed polymer of any of the polymers of claims 38 to 45.

47. The method according to any of the preceding claims, wherein the target polynucleotide has a length of 8 bases to 1000 kb.

48. The method according to claim 47, wherein the length of the target polynucleotide is from 8-15 bases, from 15-30 bases, from 30 to 50 bases, from 50 to 100 bases, from 100 to 200 bases, from 200 to 300 bases, from 300 to 500 bases, from 500 to 750 bases, from 750 to 1000 bases, from 1000 to 1500 bases, from 1500 to 3000 bases, from 3000 to 5000 bases, from 5000 to 10000 bases, from 10000 to 15000 bases, from 15000 to 20000 bases, from 20000 to 25000, from 25000 to 30000 bases, from 30000 to 35000 bases, from 35000 to 40000 bases, from 40000 to 45000 bases, from 45000 to 50000 bases, from 50000 to 75000 bases, from 75000 to 100000 bases, from 100 kb to 250 kb, from 250 to 500 kb, from 500 kb to 750 kb, from 750 kb to 1000 kb, or more than 1000 kb.

49. The method according to any of the preceding claims, wherein the length of the overlap between the probe and target polynucleotide is at least 5 nucleotides, more preferably at least 6 nuceotides, such as at least 7 nucleotides, for example 8 nucleotides, such as at least 9 nucleotides, for example at least 10 nucleotides, such as at least 15 nucleotides, for example at least 20 nucleotides, for example at least 25 nucleotides, such as at least 50 nucleotides, for example at least 100 nucleotides.

57. The method according to any of the preceding claims, wherein the label is bound to the nucleotide being complementary to the polymorphic site.

58. The method according to any of the preceding claims, wherein the label is bound to a nucleotide at least 1 nucleotide upstream or downstream to the nucleotide complementary to the polymorphic site, such as at least 2 nucleotides upstream or downstream, for example at least 3 nucleotides, such as at least 5 nucleotides, for example at least 10 nucleotides.

59. The method according to any of the preceding claims, wherein at least one probe has at least two stretches of complementarity to at least one target polynucleotide, such as at least 3 stretches, for example at least 4 stretches, such as at least 5 stretches.

60. The method according to claim 59, wherein two stretches are separated by a nucleotide sequences, which does not hybridise to the target polynucleotide.

61. The method according to any of the preceding claims, further comprising amplification of a polynucleotide prior to hybridisation.

62. The method according to claim 61, wherein the amplification comprises PCR, long range PCR, and any variant of PCR amplification.

63. The method according to claim 61, wherein the amplification comprises ligase chain reaction, asymmetric amplification, single-strand amplification, T7 amplification, NASBA (Nucleic Acid Sequence-Based Amplification), strand displacement amplification, or rolling circle amplification, or T7 polymerase amplification.

64. The method according to claim 61, wherein the amplification comprises amplification in bacteria, yeast, other cells, YAC amplification, BAC amplification or other artificial chromosome based amplifications.

71. The method according to any of the preceding claims, wherein hybridisation is carried out in solution.

72. The method according to any of the preceding claims, wherein the target or the probe is linked to a solid support prior to hybridisation.

73. The method according to claim 72, wherein said solid support comprises beads such as magnetic beads and/or the surface of a well.

74. The method according to any of the preceding claims, wherein at least one probe hybridises only to one target polynucleotide.

75. The method according to any of the preceding claims, wherein at least one probe hybridises to both a wild-type target polynucleotide and to a target polynucleotide carrying a mutation or polymorphism.

76. The method according to any of the preceding claims, wherein the at least one detectable label comprises a fluorescent label.

77. The method according to claim 76, wherein the label is selected from the list in table 2 and 3.

78. The method according to any of the preceding claims, wherein the at least one label comprises a phosphorescent label.

79. The method according to any of the preceding claims, wherein the at least one label comprises a chromogenic label such as TMB (3,3',5,5-tetramethylbenzidine).

80. The method according to any of the preceding claims, wherein recording spectral data comprises detection of signal for at least 10 discrete wavelengths, more preferably at least 20 discrete wavelengths, more preferably at least 50 discrete wavelengths, more preferably at least 100 discrete wavelengths, such as at least 200 discrete wavelengths, for example at least 250 discrete wavelengths, such as at least 300 discrete wavelengths, for example at least

88. The method according to claim 87, wherein multivariate analysis comprises general multivariate analysis, principal component analysis and extensions of this, exploratory and confirmatory factor analysis in its various forms, Cluster and latent class analysis including scaled latent class analysis, structural equation analysis, Fixed mixture analysis and combinations hereof.

89. The method according to any of the preceding claims, wherein data are treated using a neural network.

90. The method according to any of the preceding claims, wherein spectral data are recorded from hybrid polynucleotides in solution.

91. The method according to claim 90, wherein the spectral data are recorded from a solution comprising both the hybrid polynucleotide and unhybridised probe.

92. The method according to any of the preceding claims 1 to 88, wherein spectral data are recorded from hybrid polynucleotides bound to a solid support.

93. The method according to claim 92, wherein the solid support comprises a solid surface capable of immobilising a capture probe, a capture probe capable of immobilising the target polynucleotide, and a labelled detection probe capable of hybridising to the immobilised target polynucleotide.

94. The method according to claim 92, wherein the solid support is a disposable or reusable device such as but not exclusively a flow-through system.

95. The method according to claim 92, wherein the capture probe is immobilised a priori to the solid surface.

96. The method according to claim 92, wherein the capture probe is hybridised to the target before immobilisation on a solid support.

97. The method according to claim 92, wherein the capture probe(s) is/are (an) allele specific probe(s).

106. The kit according to claim 101, being in the form of a tube container with at least one probe linked to the inner surface, being a solid surface, the tube wall allowing electromagnetic radiation to pass the walls.

107. The kit according to claim 106, wherein the tube comprises more than one probe linked to more than one location, the locations being spatially separate.

108. The kit according to claim 106, wherein the tube comprises more than one probe the probes having detectably different labels.

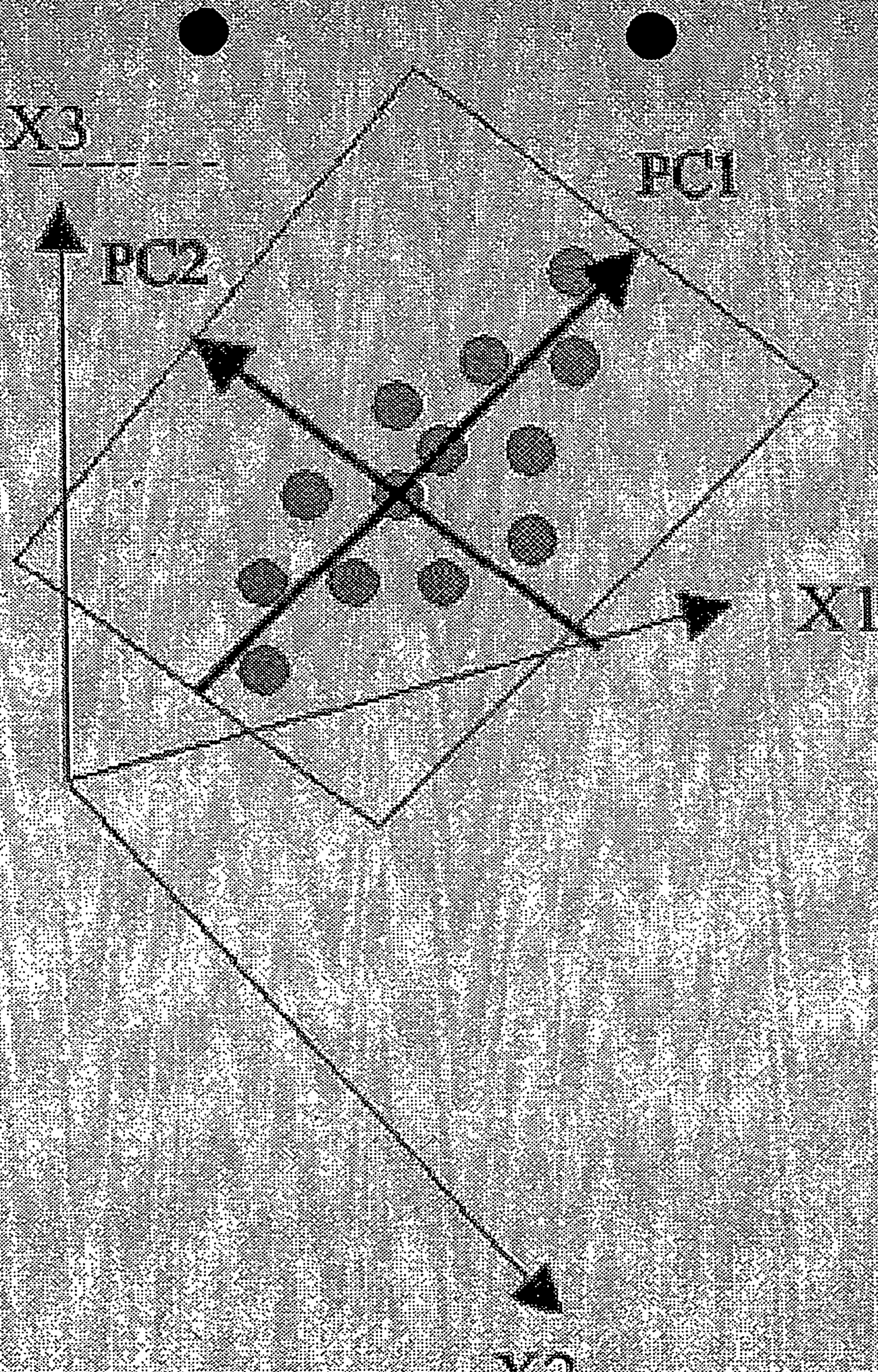
109. A system for detection of a hybrid polynucleotide comprising

i) at least one oligonucleotide probe being at least partly complementary to a target polynucleotide, the probe comprising a detectable label,
ii) a sample chamber from which electromagnetic radiation can be recorded,
iii) a source of spectrally resolved electromagnetic radiation,
iv) means for sensing and recording a spectrum of electromagnetic radiation from the sample chamber,
v) a computer unit for storing spectral data of electromagnetic radiation and having instructions to treat the recorded spectral data using multivariate analysis.

110. The system according to claim 109, further comprising a computer controlled robot to transfer solutions to the sample chamber.

111. The system according to claim 109, further comprising means to control the temperature of the sample chamber.

112. The system according to claim 109, wherein the sample chamber is in the form of a tube with at least one probe linked to the inner surface.


X3
PC1
PC2
X1
X2

Modtaget
13 SEP. 2002
PVS

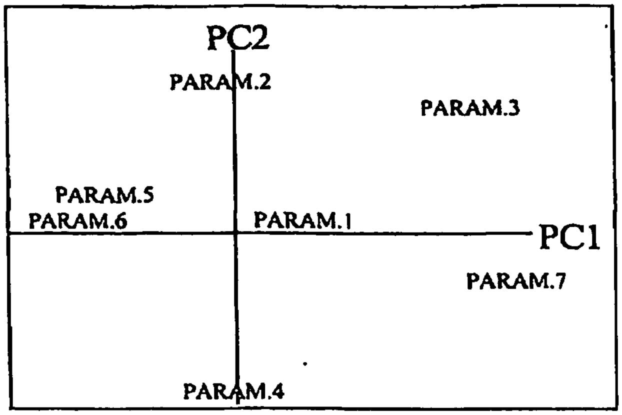


Fig. 3

Modtaget
13 SEP. 2002
PVS

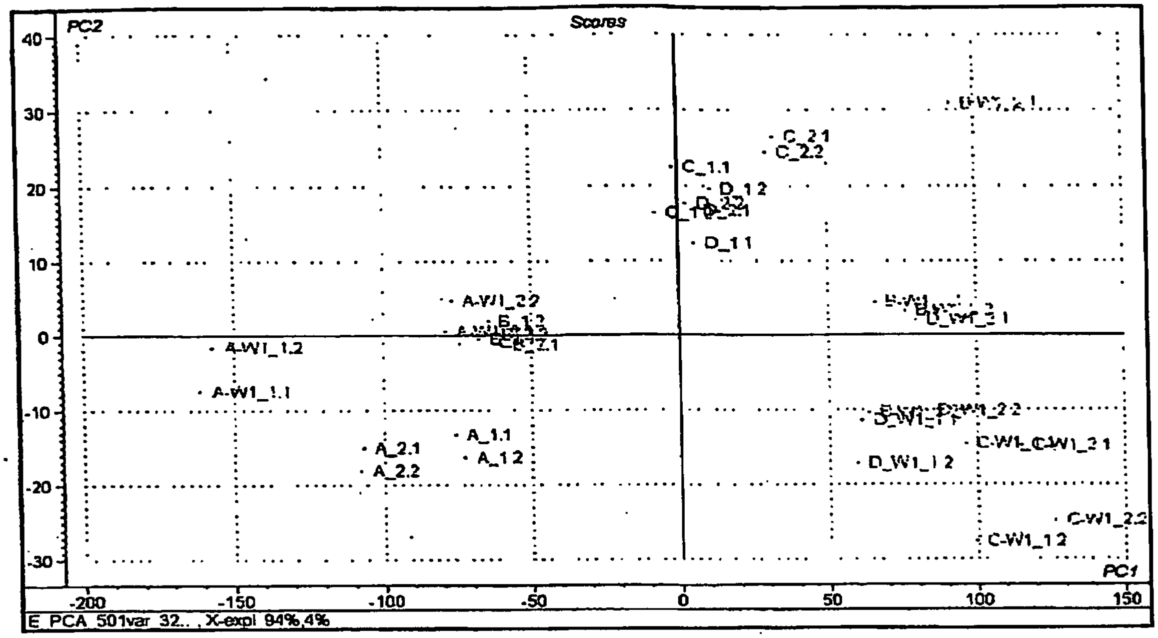


Fig. 5

Modtaget
13 SEP. 2002
PVS

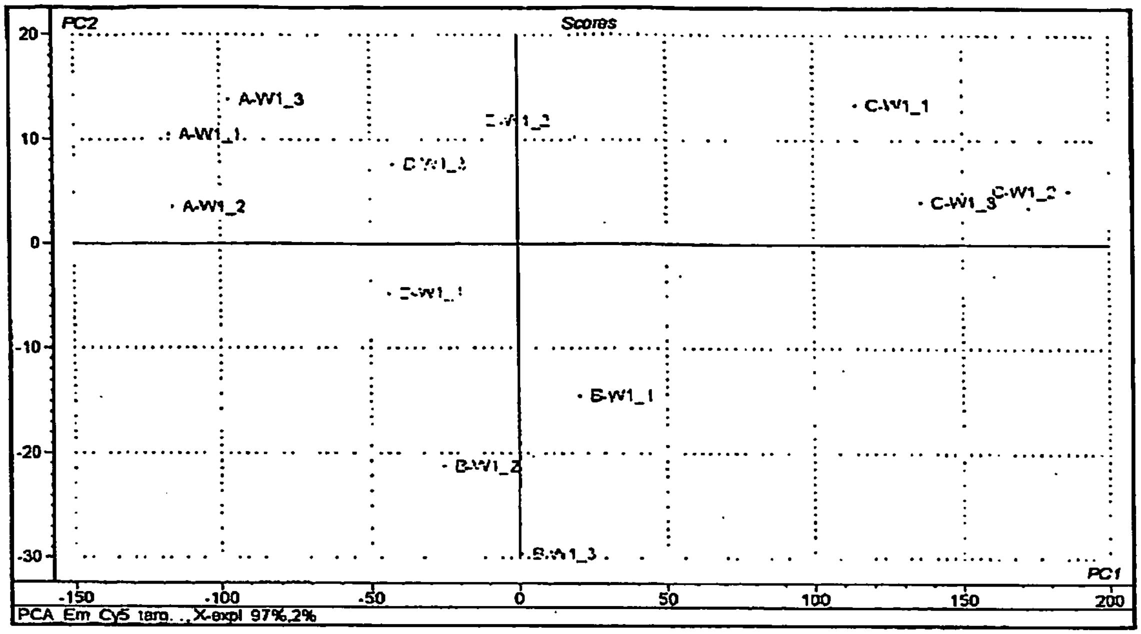


Fig. 7

Modtaget
03 SEP. 2002
PVS

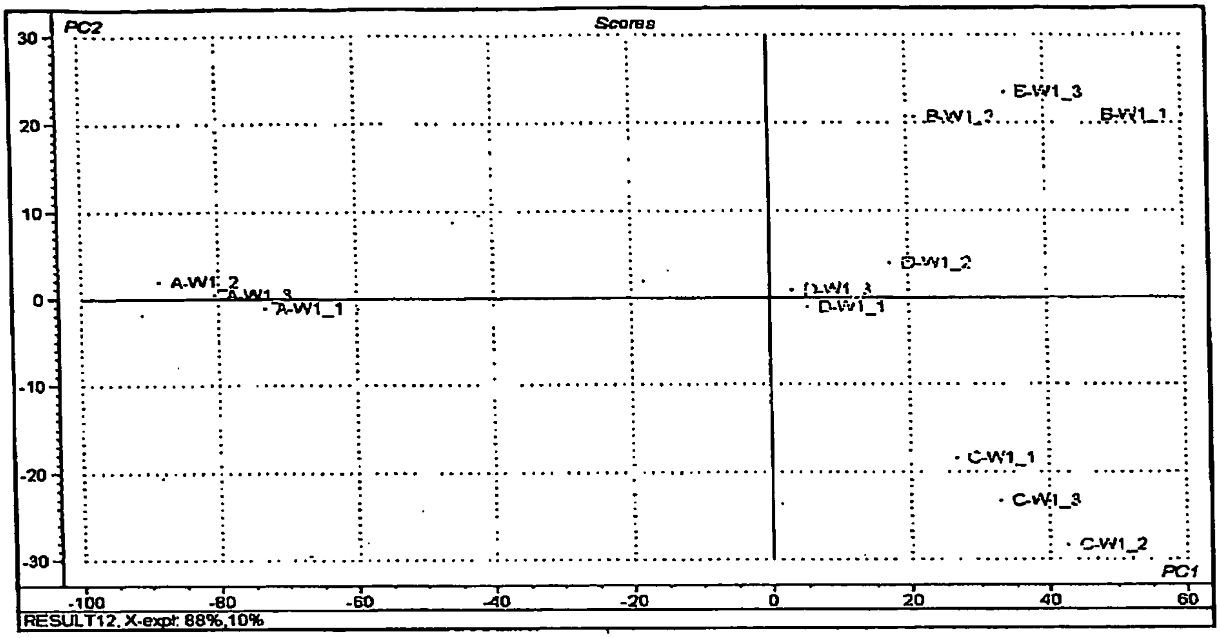


Fig. 9.

Modtaget
13 SEP. 2002
PVS

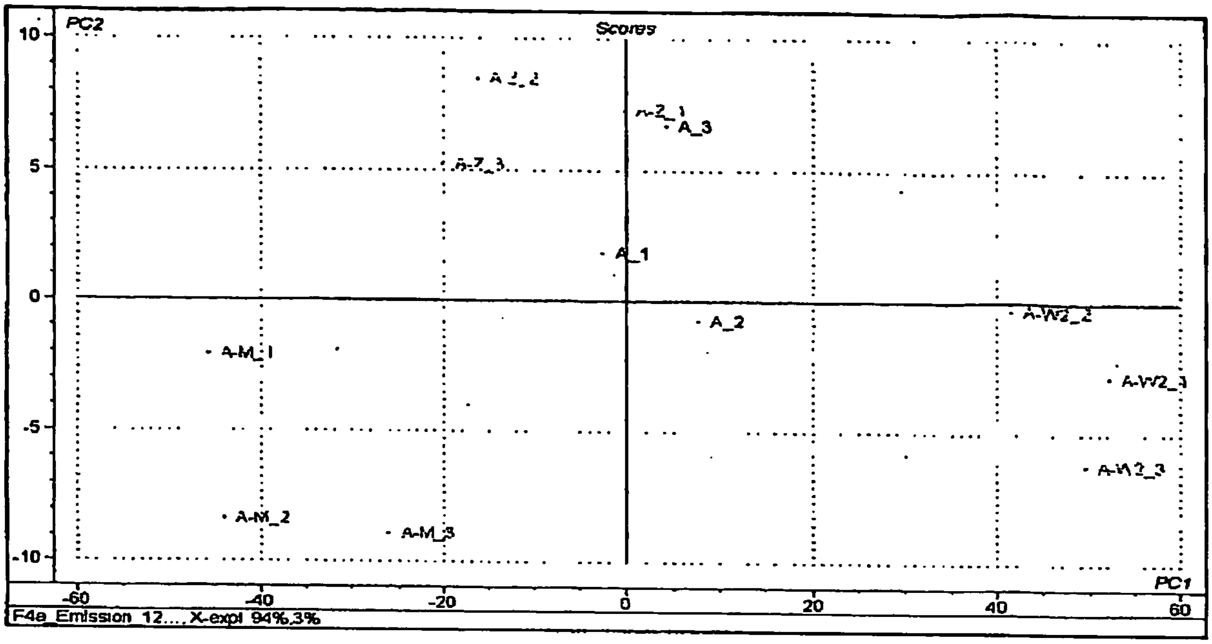


Fig. 11.

Modtaget

13 SEP. 20[illegible]

PVS

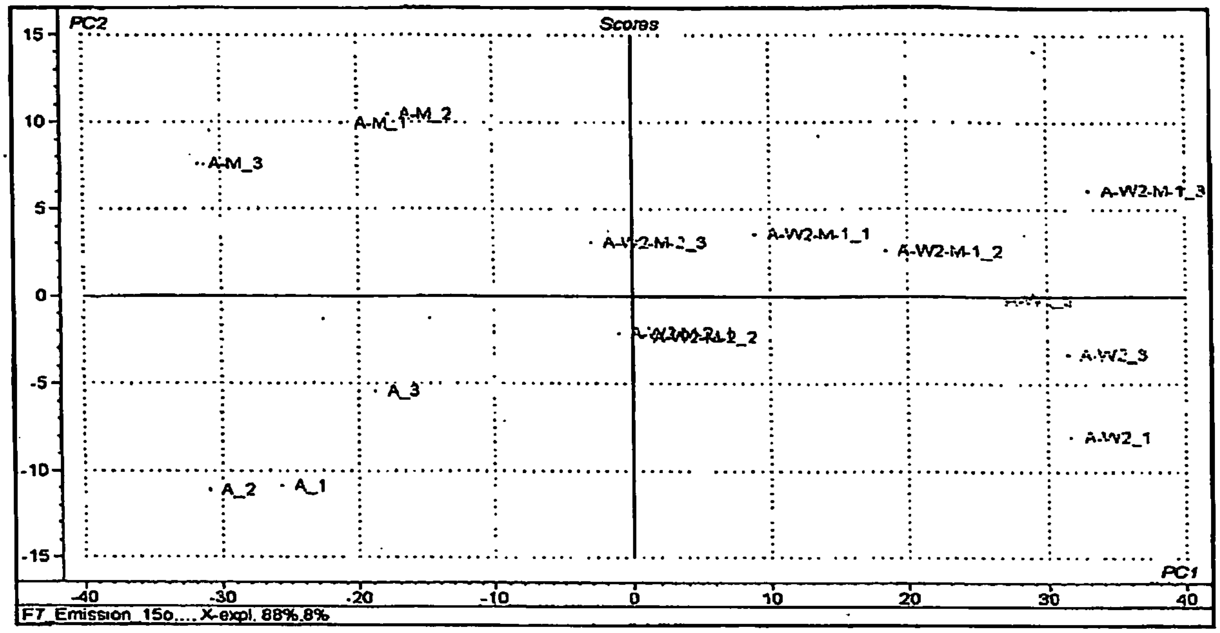


Fig. 13.

Modtaget

13 SEP. 2002

PVS

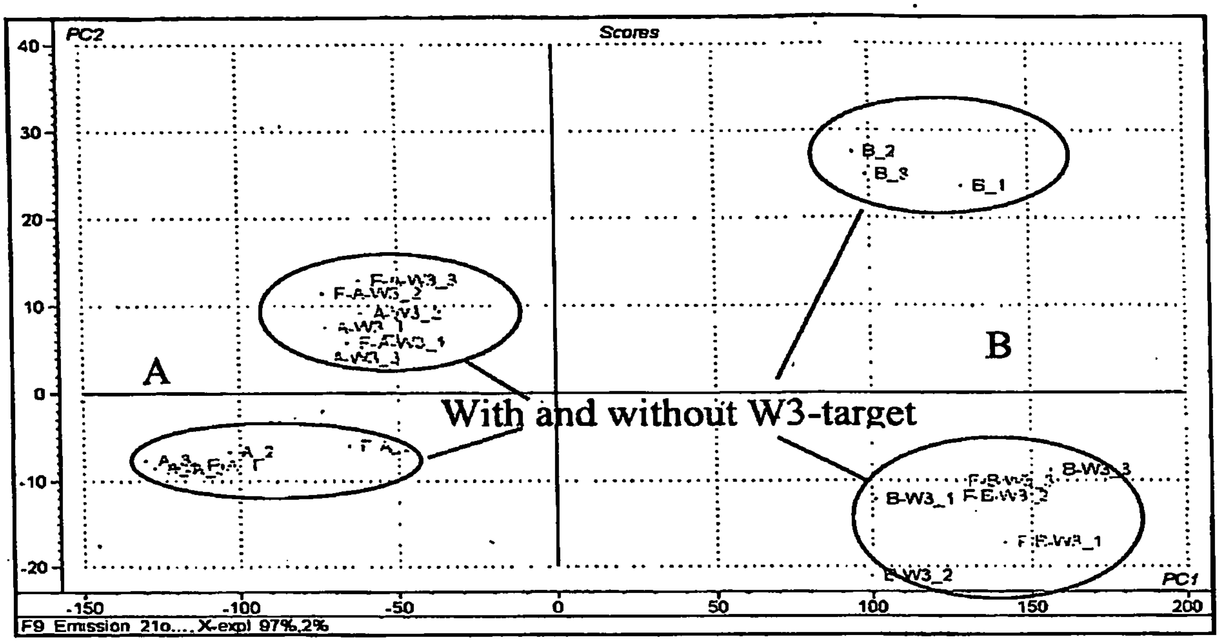


Fig. 15

Modtaget

13 SEP. 2002

PVS

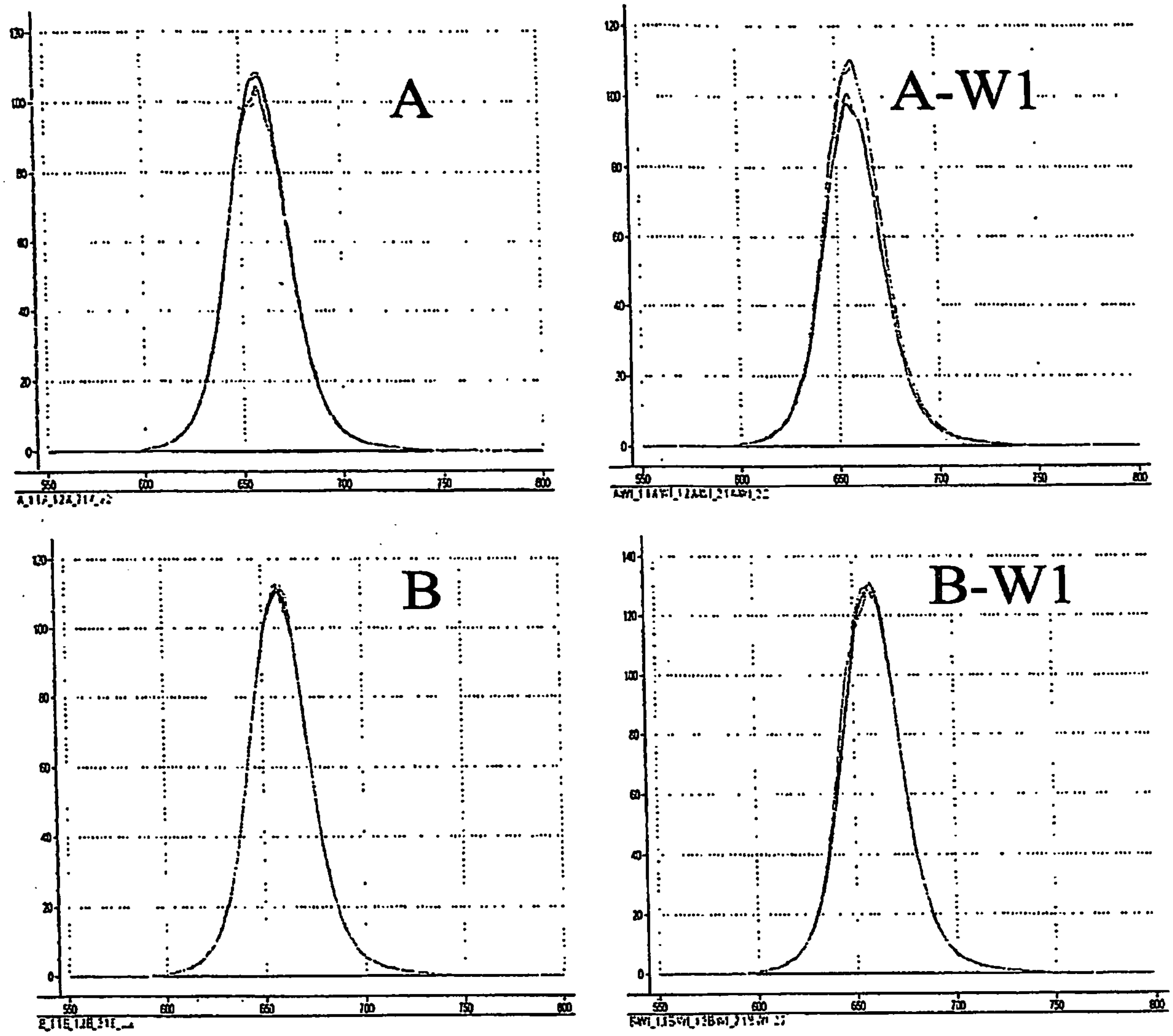


Fig. 17

13 SEP. 2002
PVS

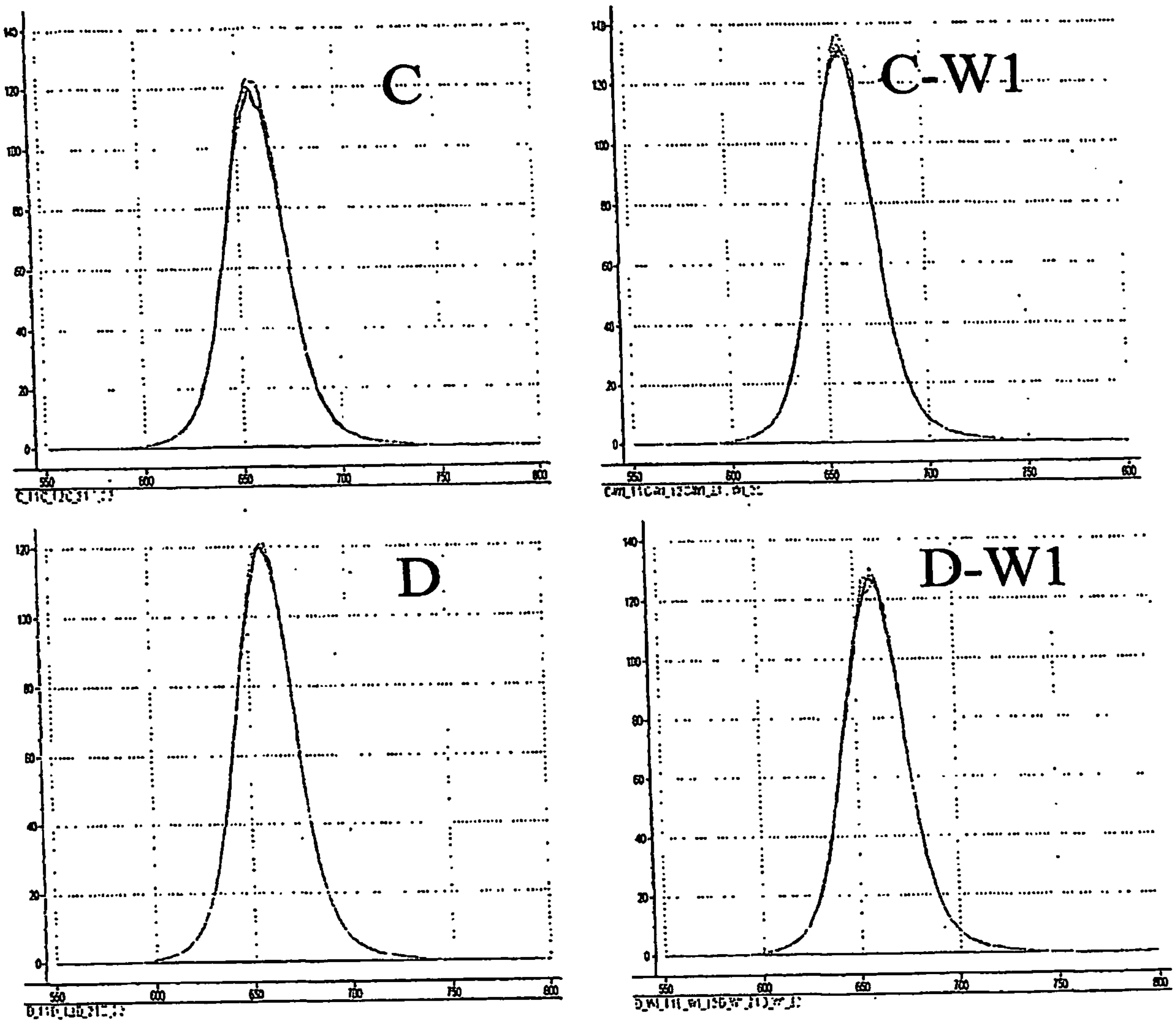


Fig. 18

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.